

機関番号：10101

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500165

研究課題名(和文) 楽譜情報に基づく記号論的音楽情報検索のための基盤技術開発

研究課題名(英文) A Study on Algorithms for Symbolic Music Information Retrieval

研究代表者

大久保 好章 (OKUBO, Yoshiaki)

北海道大学・情報科学研究科・助教

研究者番号：40271639

交付決定額(研究期間全体)：(直接経費) 3,800,000円、(間接経費) 1,140,000円

研究成果の概要(和文)：音楽情報検索のためのクラスタ抽出手法の開発を目的として、形式概念に基づくクラスタ抽出手法の拡張、および、グラフの高速クリーク探索手法の拡張を様々な角度から試みた。主な成果として、チャンスパターン(概念)の抽出、長大内包パターンの高速抽出、クラスタ構造の変化検出、真極大 k-plex の高速抽出、真極大 k-plex の繰返し探索による最大 k-plex の高速抽出が可能なアルゴリズムを得た。

研究成果の概要(英文)：The purpose of this research is to investigate fundamental methods for extracting interesting and useful clusters for Music Information Retrieval. The algorithms we have developed for this research are as follows: 1) Algorithm for Finding Top-N Chance Patterns with KeyGraph(R)-Based Importance, 2) Algorithm for Finding Top-N Colossal Patterns Based on Clique Search with Dynamic Update of Graph, 3) Algorithm for Detecting Structural Change Pattern Based on Constrained Maximal k-Plex Search, 4) Algorithm for Enumerating Maximal k-Plexes with Iterative Proper l-Plex Search and 5) Algorithm for Detecting Maximum k-Plex with Iterative Proper l-Plex Search.

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：クラスタ アルゴリズム グラフ k-plex 形式概念

## 1. 研究開始当初の背景

近年の情報通信技術の急速な発展がもたらした多種多様かつ膨大なデータに対して、我々は日常的に情報検索を行い、有用な情報を得ている。検索の対象は文書や画像など様々であるが、音楽(楽曲)もまた重要な対象のひとつであり、特にその場合は音楽情報検索と呼ばれる。

一方、大規模データから有用な情報・知識を効率良く抽出することを目的としたデータマイニングの研究が1990年代に始まったが、これも広い意味で高度な情報検索と捉えても強ち誤りではないであろう。ただし、医療や生命情報科学(バイオ)等の分野における専門的・科学的知識の発見・抽出を目指す試みは盛んに行なわれてものの、音楽データを対象としたマイニング研究は驚くべき程少ないのが現実である。

## 2. 研究の目的

代表者はこれまで、データマイニング研究の立場から、特に記号論的アプローチによるクラスタ抽出の研究を行ってきた。音楽は芸術の一分野であることからわかる通り、楽曲という対象は極めて感覚的かつ主観的に捉えられるものであり、そもそも計算機による扱いに馴染まない様にも思えるが、一方で、楽曲には楽譜という極めて離散的かつ抽象的な表現手段が存在する。古くから今もなお楽曲の抽象表現として楽譜が不可欠であるという事実は、『楽譜には楽曲の本質が表現されている』ことの確固たる裏付けであり、そこに記述された記号レベルのデータとして楽曲を扱える点は、大規模性が重視される計算処理においてはむしろ好都合であり、それにより人間の感性的な側面にどこまで切り込めるのかを見極めることは大変興味深い。そこで、本研究課題では、これまで得た知見を音楽情報検索の立場から捉え直すことにより、音楽情報検索のための基盤アルゴリズムの開発を試みる。

## 3. 研究の方法

研究は主に、代表者がこれまでに行ってきた形式概念に基づくクラスタ抽出手法を基礎として、そのさらなる拡張を試みる。具体的には、楽譜から抽出可能な音に関する情報を離散属性として表現し、通常のタグ情報(属性)とともに記号論的な処理により統一的に扱う枠組みを開発する。これにより、抽出される各クラスタは、音に限定されない、あるいは、タグ情報に限定されない複合的な属性群を内包とするものとなり、多様な解釈が与える様々なクラスタ抽出が可能となる。

これに加え、グラフ理論的な立場からも、頂点クラスタの抽出アルゴリズムの開発を行なう。近年のグラフアルゴリズムの発展により、大規模グラフからの高速なクリーク抽出が可能となった。クリークは最も理想的なクラスタを形成することから、既存の研究成果を拡張し、さらに発展させることで、より現実的かつ興味深いクラスタ抽出を可能とする枠組みを開発する。

## 4. 研究成果

研究期間において得られた主な成果は以下の通りである。

**楽曲データベースの整備：**1,000,000の楽曲から成る Million Song Dataset のデータ整形・加工処理を行い、本研究で開発したシステムで利用できる形式に整備した。具体的には、オリジナルデータとして公開されているクロマベクトル(実数ベクトル)列から、強成分のみを抽出したバイナリベクトル列へ変換することで離散化を行った。これをコード進行の代替物と捉えることで、楽曲の(ある種の)構造が明らかとなる。音楽に特化した SNS により提供された楽曲タグ情報も一部取り込んだことから、複合的視点からの楽曲間類似性も扱える様になった。

**チャンスパターン(概念)の抽出アルゴリズム設計：**意味的に離れた概念間にまたがるレアな概念をチャンスパターンと呼び、そ

の高速抽出アルゴリズムの設計を行った。特にここでは、文書から重要な (レアな) キーワードを抽出する枠組みである Key-Graph(R) におけるアイデアをもとに、”頻出パターンと共起するレアパターンは重要である”との考えを反映した制約を満たすパターンのみを高速に抽出する基本アルゴリズムを与えた。これにより、異なる視点にまたがるチャンスパターンの抽出が原理的に可能となり、多視点検索へ向けての重要な足掛かりを築いた。

### 長大内包パターンの高速抽出アルゴリズム

**設計:** 伝統的な頻出パターンマイニングシステムでは抽出が困難な、長大パターン (概念内包) の抽出アルゴリズムの設計を行った。具体的には、固定長の頻出パターン群の情報を保持するパターングラフをもとに、アイテム間の共起グラフを動的に更新しながら、長さが上位  $N$  である長大パターンを高速に列挙するアルゴリズムを設計・実装し、ベンチマークデータに対してその有効性を確認した。これにより、様々な属性 (特徴) を同時に有するパターンの抽出が可能となり、多視点検索における重要な基盤が得られた。

**クラスタ構造の変化検出手法の開発:** 様々な視点からのクラスタ構造の変化を検出するための基盤アルゴリズムの設計を行った。対象間の関係がグラフ構造で表現された場合、そこでのクラスタとは、密な部分グラフを構成する頂点集合を意味する。ここでは、所与のふたつのグラフにおける構造変化検出を、一方では擬似的な独立集合、他方では擬似的なクリークを形成する頂点集合の抽出処理と捉え、制約付き極大  $k$ -plex 探索問題としてこれを定式化し、その高速アルゴリズムを与えた。これにより、例えば、ジャンル A の楽曲ではよく用いられるが、その他のジャンルではほとんど用いられないコード進行パターンの抽出が可能となる。

**真極大  $k$ -plex の高速抽出:** 従来の  $k$ -plex 抽出法は、緩和の度合いを制御するパラメータ  $k$  が大きくなると、解を得ることが事実上不可能となる。その主要因は、緩和モデルとして不適当な膨大な  $k$ -plex の存在にある。ここでは真  $k$ -plex の概念を新たに導入し、抽出対象をサイズ下限制約を満たす真  $k$ -plex に限定することで、連結性と 50% の密度を保証する妥当な  $k$ -plex のみを高速に抽出する。探索過程では、頂点間の上限距離に基づく分枝の抑制、および、真  $k$ -plex 性に基づく枝刈りが可能であり、結果として既存手法では解を得ることができなかった  $k$  に対しても高速に動作することが確認できた。

### 真極大 $k$ -plex の繰返し探索による最大 $k$ -plex の高速抽出

**真極大  $k$ -plex 抽出法** をもとに、最大  $k$ -plex 抽出問題においても、より大きな  $k$  に対応できる手法を開発した。具体的には、緩和パラメータ  $k$  を段階的に強化しながら真極大  $k$ -plex 探索を繰返すことで、最大  $k$ -plex を高速に抽出するアルゴリズムを与えた。繰返しの各段階においては、既に得られている暫定最大  $k$ -plex のサイズを用いた探索の分枝限定が可能であるが、初期の段階で大きなサイズの  $k$ -plex が見つかる傾向にあることから、繰返しが進むにつれて分枝限定効果はより強力となる。ベンチマークグラフによる実験では、既存の最速アルゴリズムと比較して 100 倍から 1000 倍の高速化を観察した。

また、クラスタ構造の変化検出手法における抽出対象は  $k$ -plex であることから、本アルゴリズムを用いることで、構造変化検出をさらに高速化することも期待できる。

## 5. 主な発表論文等

### 【雑誌論文 (計 10 件)】

1. [Yoshiaki Okubo, Makoto Haraguchi and Takeshi Nakajima: Finding Rare](#)

- Patterns with Weak-Correlation Constraint: Progress in Indicative and Chance Patterns, Advances in Chance Discovery - Extended Selection from International Workshops, Yukio Ohsawa and Akinori Abe (eds.), ISBN: 978-3-642-30113-1, 250 pages, Studies in Computational Intelligence, 423, pp. 231 - 246, Springer, 査読有, 2013.
2. Yoshiaki Okubo, Makoto Haraguchi and Etsuji Tomita: Relational Change Pattern Mining Based on Modularity Difference, Proceedings of the 7th Multi-Disciplinary International Workshop on Artificial Intelligence - MIWAI'13, Lecture Notes in Artificial Intelligence, Springer-LNAI 8271, pp. 187 - 198, 査読有, 2013.
  3. Hongjie Zhai, Makoto Haraguchi, Yoshiaki Okubo, Kiyota Hashimoto and Sachio Hirokawa: Shifting Concepts to Their Associative Concepts via Bridges, Proceedings of the 9th International Conference on Machine Learning and Data Mining in Pattern Recognition - MLDM'13, Lecture Notes in Artificial Intelligence, Springer-LNAI 7988, pp. 586 - 600, 査読有, 2013.
  4. Yoshiaki Okubo and Makoto Haraguchi: Finding Top-N Colossal Patterns Based on Clique Search with Dynamic Update of Graph, Proceedings of The 10th International Conference on Formal Concept Analysis - ICFCA'12, Lecture Notes in Artificial Intelligence, Springer-LNAI 7278, pp. 244 - 259, 査読有, 2012.
  5. Yoshiaki Okubo, Makoto Haraguchi and Etsuji Tomita: Structural Change Pattern Mining Based on Constrained Maximal  $k$ -Plex Search, Proceedings of the 15th International Conference on Discovery Science - DS'12, Lecture Notes in Artificial Intelligence, Springer-LNAI 7569, pp. 284 - 298, 査読有, 2012.
  6. Li Aixiang, Makoto Haraguchi and Yoshiaki Okubo: Top-N Minimization Approach for Indicative Correlation Change Mining, Proceedings of the 8th International Conference on Machine Learning and Data Mining in Pattern Recognition - MLDM'12, Lecture Notes in Artificial Intelligence, Springer-LNAI 7376, pp. 102 - 116, 査読有, 2012.
  7. Li Aixiang, Makoto Haraguchi, Yoshiaki Okubo and Etsuji Tomita: Finding What Changes for Two Graphs Constructed from Different Time Intervals, Proceedings of the 2012 IIAI International Conference on Advanced Applied Informatics - IIAI AAI'12, pp. 48 - 53, 査読有, 2012.
  8. Li Aixiang, Makoto Haraguchi and Yoshiaki Okubo: Contrasting Correlations by an Efficient Double-Clique Condition, Transactions on Machine Learning and Data Mining, 5(1), pp. 3 - 22, ibai Publishing, 査読有, 2012.
  9. Yoshiaki Okubo, Makoto Haraguchi and Sachio Hirokawa: Finding Top-N Chance Patterns with KeyGraph(R)-Based Importance, Proceedings of The 15th International Conference on Knowledge-Based and Intelligent Information &

Engineering Systems - KES'11 (Part II), Lecture Notes in Artificial Intelligence, Springer-LNAI 6882, pp. 457 - 468, 査読有, 2011.

10. Li Aixiang, Makoto Haraguchi and Yoshiaki Okubo: Contrasting Correlations by an Efficient Double-Clique Condition, Proceedings of the 7th International Conference on Machine Learning and Data Mining in Pattern Recognition - MLDM'11, Lecture Notes in Artificial Intelligence, Springer-LNAI 6871, pp. 469 - 483, 査読有, 2011.

### 【学会発表 (計 11 件)】

1. 大久保 好章・松平 将宜・原口 誠: 反復真極大 k-Plex 探索による最大 k-Plex の高速抽出, 人工知能学会研究会資料, SIG-FPAI-B304, pp. 53 - 60, 宮古島マリンターミナル (沖縄県宮古島市), 2014 年 3 月 8 日.
2. 大久保 好章・原口 誠: 真極大 k-Plex の高速抽出に関する一考察, 人工知能学会研究会資料, SIG-FPAI-B302, pp. 31 - 38, 愛媛大学・情報メディアセンター (愛媛県松山市), 2013 年 11 月 28 日.
3. 大久保 好章・原口 誠: 密度を考慮した極大 k-Plex 抽出に関する一考察, 人工知能学会研究会資料, SIG-FPAI-B301, pp. 31 - 36, 稚内日口友好会館 (北海道稚内市), 2013 年 7 月 18 日.
4. Gupei Qin, Makoto Haraguchi, Yoshiaki Okubo: Mining Contrast Concepts Representing Differences, 人工知能学会全国大会 (第 27 回) 論文集, 1C3-5, 富山国際会議場 (富山県富山市), 2013 年 6 月 4 日.
5. Hongjie Zhai, Makoto Haraguchi, Yoshiaki Okubo, Kiyota Hashimoto

and Sachio Hirokawa: A Framework of Associative Search by Mediators, 人工知能学会全国大会 (第 27 回) 論文集, 1F5-6, 富山国際会議場 (富山県富山市), 2013 年 6 月 4 日.

6. Hongjie Zhai, Mingyu Bai, Makoto Haraguchi, Yoshiaki Okubo, Kiyota Hashimoto and Sachio Hirokawa: A Framework of Association Search by Mediators (人を介する連想検索の提案), 人工知能学会研究会資料, SIG-FPAI-B203, pp. 113 - 118, 石垣市民会館 (沖縄県石垣市), 2013 年 1 月 25 日.
7. 大久保 好章・原口 誠: 動的パターングラフを用いた Top-N 長大パターン抽出, 人工知能学会研究会資料, SIG-FPAI-B202, pp. 55 - 60, 慶応義塾大学・來往舎 (神奈川県横浜市), 2012 年 11 月 17 日.
8. 大久保 好章・原口 誠: 擬似クリークと負の制約を用いたグラフの構造変化検出, 人工知能学会全国大会 (第 26 回) 論文集, 4C1-R-6-7, 山口県教育会館・他 (山口県山口市), 2012 年 6 月 15 日.
9. エラウィンディ サラ・原口 誠・大久保 好章・富田 悦次: クリーク全列挙に基づく構造変化検出アルゴリズム, 情報処理学会研究報告, Vol. 2011-MPS-087 No. 32, 指宿市市民会館 (鹿児島県指宿市), 2012 年 3 月 2 日.
10. Yoshiaki Okubo and Makoto Haraguchi: An Algorithm for Finding Indicative Concepts Connecting Larger Concepts Based on Structural Constraints, Contributions to ICFCA 2011, The 9th International Conference on Formal Concept Analysis - ICFCA'11, pp. 53 - 68, Univer-

sity of Nicosia (Nicosia, Cyprus), 4th  
May, 2011.

11. 原口 誠・大久保 好章・富田 悦次・吉  
岡真治：変化検出のための極大整合  
連結集合，人工知能学会研究会資料，  
SIG-FPAI-B103, pp. 43 - 48, 慶応  
義塾大学・來往舎 (神奈川県横浜市)，  
2011 年 12 月 16 日.

## 6. 研究組織

### (1) 研究代表者

大久保 好章 (Yoshiaki OKUBO)

北海道大学・大学院情報科学研究科・

助教

研究者番号：40271639